

# Installation guide

# **Table of content**

1. Project overview
2. Required folder structure
3. Deliverables and installations instructions
4. Required setup files
5. Required setup tables

# **Project overview**

## *(Data load from sources to staging database)*

The project goal is to load staging database from various data sources which includes ad configuration data, ad performance data, lead data etc. In first two sprints ETL scripts which will be fetching data from APIs and loading to database will be provided along with postgres schema and setup data required for the same. In this document, we have divided installation instructions source/deliverable wise:

- 1) Face book code – The code which will fetch data from face book API and loads into staging DB. There are separate batches for configuration and performance data fetch from face book.
- 2) Cake code – The code which is required to fetch leads by affiliates, leads by buyer, cake post data from Cake API and loads into staging DB. There is one batch for all three kinds of data fetch processes.
- 3) Keen code – The code which fetches data from keen API and loads data to staging DB. There is one batch which is running all the sub processes for this section.
- 4) Ad words code – The code which fetches data from ad words API and loads into staging DB. There is 1 batch for all sub processes running for this section.
- 5) Bing code – The code which fetches data from keen API and loads data to staging DB. There are 2 separate batches in this section, one for configuration data and one for performance data.

# **Required folder structure**

Currently on test VM everything resides under /home/ubuntu.

Under this directly following directory needs to be created:

- stage – This directory will contain temporary files created during jobs run.
- config\_files – This directory will contain configuration files for each source. These files will contain target database credentials.
- target – This directory contains data files for some of the sources.
- logs – This directory will contain job log for each run.

Same hierarchy needs to be reflected.

Note: If path needs to be changed to something else than /home/ubuntu, same should be reflected in table entries for config\_table.

# Deliverables and installations instructions

All the ETL scripts have been placed at following path:  
/home/ubuntu/deliverables/etl\_scripts

## I. Deliverable name: FETCH\_DATA\_PER\_ACCOUNT.zip (Facebook config data processing batch)

Description: This archive contains script which will fetch configuration data from facebook for all accounts. The script name is FETCH\_DATA\_PER\_ACCOUNT\_run.sh

Script dependencies – Script requires config\_table to be loaded with required data.

Also configuration file containing database parameter values should be present at following path: /home/ubuntu/config\_files/facebook\_config.txt

Follow following steps for installation:

- a. Create an entry in table jobs such as:

```
INSERT INTO jobs ("FromEmailID", "JobName", "JobDescription", "JobSuccessEmailID",  
"JobFailureEmailID", "CreatedDate", "UpdatedDate", "IsActive", "JobSuccess_CC_EmailID",  
"JobFailure_CC_EmailID", "JobGroup", "JobID") VALUES ('cybytetech@yahoo.com',  
'FETCH_DATA_PER_ACCOUNT', NULL, 'poonamt@cybytetech.com',  
'poonamt@cybytetech.com', NULL, NULL, 1, 'poonamt@cybytetech.com',  
'poonamt@cybytetech.com', NULL, 6);
```

**Note: Please put required email addresses where you want respective email notification to be delivered. If you want success email to be delivered to multiple email accounts, separate them by a comma. Likewise for other email lists.**

- b. Unzip FETCH\_DATA\_PER\_ACCOUNT.zip on server.
- c. Please make sure facebook\_config.txt at following path: /home/ubuntu/config\_files

Also config files should have following entries:

```
DB_HOST=<Database host name>  
DB_NAME=<Database name>  
DB_PORT=<Database Port>  
DB_SCHEMA=<Database Schema>  
DB_USER=<Database User>  
DB_PASSWORD=<Database Password>
```

- d. Please make sure following entries are present in table config\_table:

```
INSERT INTO config_table (jobname, key, value) VALUES ('api_facebook', 'ACCESS_TOKEN', <Please provide facebook  
access token here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'LOG_DIR_PATH', '/home/ubuntu/logs');
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'TARGET_DIR_PATH', '/home/ubuntu/target');

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'STAGING_DIR_PATH', '/home/ubuntu/stage');

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'SMTP_PORT', <Please provide SMTP port here>);

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'SMTP_HOST', <Please provide SMTP host here>);

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'USER_NAME', <Please provide SMTP username here>);

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'USER_PASSWORD', <Please provide SMTP password here>);
```

- e. Please make sure facebook\_config.txt at following path:/home/ubuntu/config\_files
- f. Run/Schedule script FETCH\_DATA\_PER\_ACCOUNT /FETCH\_DATA\_PER\_ACCOUNT\_run.sh (.bat file in same folder for windows server) . The script file should have read and execute privileges set properly for schedule to work.
- g. To check log go to \$LOG\_DIR\_PATH/fb\_config\_process.txt
- h. For every run an entry will be made to table runlog which will have start and end timestamp for run.

## **Output data tables**

Table name	Description
fb_ad_accounts	Table containing all ad accounts
fb_campaigns	Table containing all campaigns
fb_ads	Table containing all ads
fb_ads_track_specs	Table containing all ads tracking specifications
fb_ad_creatives	Table containing all ad creatives
fb_adsets	Table containing all ad sets
fb_ad_labels	Table containing all ad labels
fb_ad_users	Table containing all ad users
fb_cust_audiences	Table containing custom audiences data
fb_conn_objects	Table containing connection objects data e.g. pages data
fb_tageting	Table containing targeting information for ads

## **II. Deliverable name: FETCH\_INSIGHTS\_ALL\_ADS.zip(Facebook performance data processing batch)**

**Description:** This archive contains script which will fetch performance data from facebook for all accounts. The script name is FETCH\_INSIGHTS\_ALL\_ADS\_run.sh

**Script dependancies –** Script requires config\_table to be loaded with required data.Also configuration file containing database parameter values should be present at following path: /home/ubuntu/config\_files/facebook\_config.txt

Follow following steps for installation:

- i. Create an entry in table jobs such as:

```
INSERT INTO jobs ("FromEmailID", "JobName", "JobDescription", "JobSuccessEmailID", "JobFailureEmailID",  
"CreatedDate", "UpdatedDate", "IsActive", "JobSuccess_CC_EmailID", "JobFailure_CC_EmailID", "JobGroup",  
"JobID") VALUES ('cybytetech@yahoo.com', 'FETCH_INSIGHTS_ALL_ADS', NULL, 'poonamt@cybytetech.com',  
'poonamt@cybytetech.com', NULL, NULL, 1, 'poonamt@cybytetech.com', 'poonamt@cybytetech.com', NULL, 6);
```

**Note:** Please put required email addresses where you want respective email notification to be delivered. If you want success email to be delivered to multiple email accounts, separate them by a comma. Likewise for other email lists.

- j. Make sure tables fb\_ads and db\_ad\_accounts are present in database. Otherwise run FETCH\_DATA\_PER\_ACCOUNT to create same.
- k. Unzip FETCH\_INSIGHTS\_ALL\_ADS.zip on server.
- l. Please make sure facebook\_config.txt at following path: /home/ubuntu/config\_files

Also config files should have following entries:

```
DB_HOST=<Database host name>  
DB_NAME=<Database name>  
DB_PORT=<Database Port>  
DB_SCHEMA=<Database Schema>  
DB_USER=<Database User>  
DB_PASSWORD=<Database Password>
```

- m. Please make sure following entries are present in table config\_table:

```
INSERT INTO config_table(jobname, key, value) VALUES ('api_facebook', 'ACCESS_TOKEN', <Please provide facebook  
access token here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'LOG_DIR_PATH', '/home/ubuntu/logs');
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'TARGET_DIR_PATH', '/home/ubuntu/target');
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'STAGING_DIR_PATH', '/home/ubuntu/stage');
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'SMTP_PORT', <Please provide SMTP port  
here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'SMTP_HOST', <Please provide SMTP host  
here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'USER_NAME', <Please provide SMTP username  
here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'USER_PASSWORD', <Please provide SMTP  
password here>);
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('api_facebook', 'DATE_RANGE',  
'&date_preset=last_14_days&time_increment=1');
```

**Note:** This is the parameter which needs to change if one wants to load data for different timeframe. Please refer following info to know what others value can be provided to this parameter:

#### `date_preset`

enum{today, yesterday, last\_3\_days, this\_week, last\_week, last\_7\_days, last\_14\_days, last\_28\_days, last\_30\_days, last\_90\_days, this\_month, last\_month, this\_quarter, last\_3\_months, lifetime}

Default value: `lifetime`

Represents a relative time range. This field is ignored if `time_range` or `time_ranges` is specified.

#### `time_increment`

enum{monthly, all\_days} or integer

Default value: `all_days`

If it is an integer, it is the number of days from 1 to 90. After you pick a reporting period by using `time_range` or `date_preset`, you may choose to have the results for the whole period, or have results for smaller time slices. If "all\_days" is used, it means one result set for the whole period. If "monthly" is used, you will get one result set for each calendar month in the given period. Or you can have one result set for each N-day period specified by this param. This param is ignored if `time_ranges` is specified.

#### `time_range`

{'since'=>YYYY-MM-DD,'until'=>YYYY-MM-DD}

A single time range object. UNIX timestamp not supported. This param is ignored if `time_ranges` is provided.

Please refer following link for more info about the same:<https://developers.facebook.com/docs/marketing-api/reference/adgroup/insights/>

- n. Run/Schedule script `FETCH_INSIGHTS_ALL_ADS / FETCH_INSIGHTS_ALL_ADS_run.sh` (.bat file in same folder for windows server) . The script file should have read and execute privileges set properly for schedule to work.
- o. To check log go to `$LOG_DIR_PATH/fb_insights_process.txt`
- p. For every run an entry will be made to table `runlog` which will have start and end timestamp for run.

### **Output data tables**

Table name	Description
<code>fb_ads_perf_data</code>	Table containing all ads performance data
<code>fb_ads_perf_data_plc_dev</code>	Table containing all ads performance data aggregated by placement and impression device
<code>fb_ad_insights_unq_actions</code>	Table containing all ads unique actions parsed from performance data



### III. Deliverable name: ALL\_ADWORDS\_CLIENTS\_LOAD.zip(Google adwords data processing batch)

Description: This archive contains script which will fetch configuration and performance data from google adwords for all accounts. The script name is ALL\_ADWORDS\_CLIENTS\_LOAD\_run.sh

Script dependencies – Script requires config\_table to be loaded with required data.

Also configuration file containing database parameter values should be present at following path: /home/ubuntu/config\_files/adwords\_config.txt

Follow following steps for installation:

- a. Create an entry in table jobs such as:

```
INSERT INTO jobs ("FromEmailID", "JobName", "JobDescription", "JobSuccessEmailID", "JobFailureEmailID", "CreatedDate", "UpdatedDate", "IsActive", "JobSuccess_CC_EmailID", "JobFailure_CC_EmailID", "JobGroup", "JobID") VALUES ('cybytetech@yahoo.com', 'ALL_ADWORDS_CLIENTS_LOAD', NULL, 'poonamt@cybytetech.com', 'poonamt@cybytetech.com', NULL, NULL, 1, 'poonamt@cybytetech.com', 'poonamt@cybytetech.com', NULL, 6);
```

**Note: Please put required email addresses where you want respective email notification to be delivered. If you want success email to be delivered to multiple email accounts, separate them by a comma. Likewise for other email lists.**

- b. Unzip ALL\_ADWORDS\_CLIENTS\_LOAD.zip on server.
- c. Please make sure adwords\_config.txt at following path:/home/ubuntu/config\_files

Also configuration files should have following entries:

```
DB_HOST=<Database host name>
DB_NAME=<Database name>
DB_PORT=<Database Port>
DB_SCHEMA=<Database Schema>
DB_USER=<Database User>
DB_PASSWORD=<Database Password>
```

- d. Please make sure following entries are present in table config\_table:

```
INSERT INTO config_table (jobname, key, value) VALUES ('api_adwords', 'KEY_FILE', '/home/ubuntu/keytest-a2fd1768d1c2.p12');
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('api_adwords', 'ACC_EMAIL', <Please provide account email associated with service account>);
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('api_adwords', 'DEV_TOKEN', <Please provide dev token here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'LOG_DIR_PATH', '/home/ubuntu/logs');
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'TARGET_DIR_PATH', '/home/ubuntu/target');
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'STAGING_DIR_PATH', '/home/ubuntu/stage');
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'SMTP_PORT', <Please provide SMTP port here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'SMTP_HOST', <Please provide SMTP host here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'USER_NAME', <Please provide SMTP username here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'USER_PASSWORD', <Please provide SMTP password here>);
```

- e. Run/Schedule script ALL\_ADWORDS\_CLIENTS\_LOAD/ALL\_ADWORDS\_CLIENTS\_LOAD\_run.sh (.bat file in same folder for windows server) . The script file should have read and execute privileges set properly for schedule to work.
- f. To check log go to \$LOG\_DIR\_PATH/adwords\_config\_process.txt
- g. For every run an entry will be made to table runlog which will have start and end timestamp for run.

### **Output data tables**

Table name	Description
adw_acc_perf_report_1	Table containing all adwords account performance report data
adw_acc_perf_report_2	Table containing all adwords account performance report data – columns which are not compatible with those in adw_acc_perf_report_1
adw_ad_perf_report_config	Table containing urls from ad_performance_report
adw_ad_performance_report_1	Table containing all ads performance data
adw_ad_performance_report_2	Table containing all ads performance data– columns which are not compatible with those in adw_acc_perf_report_1
adw_ad_performance_report_3	Table containing all ads performance data– columns which are not compatible with those in adw_acc_perf_report_1
adw_click_performance_report	Table containing all click performance data
adw_dest_url_criteria	Table containing all destination url data with a criteria
adw_dest_url	Table containing all destination url data
adw_kw_perf_report_1	Table containing all keyword performance report data
adw_kw_perf_report_2	Table containing all keyword performance report data - columns which are not compatible with those in adw_kw_perf_report_1
adw_kw_perf_report_3	Table containing all keyword performance report data - columns which are not compatible with those in adw_kw_perf_report_1 and adw_kw_perf_report_2
adw_search_query_performance_report	Table containing all search query performance report data
adw_search_query_performance_report_1	Table containing all search query performance report data - columns which are not compatible with those in adw_search_query_performance_report

#### IV. Deliverable name: MASTER\_KEEN.zip(KEEN API data processing batch)

Description: This archive contains script which will fetch keen api data. The script name is master\_keen\_events\_run.sh

Script dependencies – Script requires config\_table to be loaded with required data.

Also configuration file containing database parameter values should be present at following path: /home/ubuntu/config\_files/keen\_config.txt

Follow following steps for installation:

- a. Create an entry in table jobs such as:

```
INSERT INTO jobs ("FromEmailID", "JobName", "JobDescription", "JobSuccessEmailID", "JobFailureEmailID",
"CreateDate", "UpdatedDate", "IsActive", "JobSuccess_CC_EmailID", "JobFailure_CC_EmailID", "JobGroup",
"JobID") VALUES ('cybytetech@yahoo.com', master_keen_events', NULL, 'poonamt@cybytetech.com',
'poonamt@cybytetech.com', NULL, NULL, 1, 'poonamt@cybytetech.com', 'poonamt@cybytetech.com', NULL, 6);
```

**Note: Please put required email addresses where you want respective email notification to be delivered. If you want success email to be delivered to multiple email accounts, separate them by a common. Likewise for other email lists.**

- b. Unzip MASTER\_KEEN\_EVENTS.zip on server.
- c. Please make sure adwords\_config.txt at following path:/home/ubuntu/config\_files

Also configuration files should have following entries:

```
DB_HOST=<Database host name>
DB_NAME=<Database name>
DB_PORT=<Database Port>
DB_SCHEMA=<Database Schema>
DB_USER=<Database User>
DB_PASSWORD=<Database Password>
```

- d. Please make sure following entries are present in table config\_table:

```
INSERT INTO config_table (jobname, key, value) VALUES ('keen_api', 'PROJECT_ID', '55b221a490e4bd410e6bac59');
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('keen_api', 'READ_KEY', <Please provide API key here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'LOG_DIR_PATH', '/home/ubuntu/logs');
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'TARGET_DIR_PATH', '/home/ubuntu/target');
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'STAGING_DIR_PATH', '/home/ubuntu/stage');
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'SMTP_PORT', <Please provide SMTP port here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'SMTP_HOST', <Please provide SMTP host here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'USER_NAME', <Please provide SMTP username here>);
```

```
INSERT INTO config_table(jobname, key, value) VALUES ('common', 'USER_PASSWORD', <Please provide SMTP password here>);
```

- e. Run/Schedule script master\_keen\_events/master\_keen\_events\_run.sh (.bat file in same folder for windows server). The script file should have read and execute privileges set properly for schedule to work.
- f. To check log go to \$LOG\_DIR\_PATH/keen\_process.txt
- g. For every run an entry will be made to table runlog which will have start and end timestamp for run.

### **Output data tables**

Table name	Description
keen_event_collections	Table containing all collections data

## **V. Deliverable name: MASTER\_CAKE.zip(Cake data processing batch)**

**Description:** This archive contains script which will fetch from cake for all events. The script name is master\_cake\_run.sh

Script dependencies – Script requires config\_table to be loaded with required data. Also configuration file containing database parameter values should be present at following path: /home/ubuntu/config\_files/facebook\_config.txt

Follow following steps for installation:

- a. Create an entry in table jobs such as:

```
INSERT INTO jobs ("FromEmailID", "JobName", "JobDescription", "JobSuccessEmailID", "JobFailureEmailID", "CreatedDate", "UpdatedDate", "IsActive", "JobSuccess_CC_EmailID", "JobFailure_CC_EmailID", "JobGroup", "JobID") VALUES ('cybytetech@yahoo.com', 'master_cake', NULL, 'poonamt@cybytetech.com', 'poonamt@cybytetech.com', NULL, NULL, 1, 'poonamt@cybytetech.com', 'poonamt@cybytetech.com', NULL, 6);
```

**Note: Please put required email addresses where you want respective email notification to be delivered. If you want success email to be delivered to multiple email accounts, separate them by a comma. Likewise for other email lists.**

- b. Unzip MASTER\_CAKE.zip on server.
- c. Please make sure cake\_config.txt at following path:/home/ubuntu/config\_files

Also config files should have following entries:

DB\_HOST=<Database host name>

DB\_NAME=<Database name>

DB\_PORT=<Database Port>  
DB\_SCHEMA=<Database Schema>  
DB\_USER=<Database User>  
DB\_PASSWORD=<Database Password>

- d. Please make sure following entries are present in table config\_table:

```
INSERT INTO config_table (jobname, key, value) VALUES ('cake_api', 'GROUP_DIR', 'ASC');

INSERT INTO config_table (jobname, key, value) VALUES ('cake_api', 'BUYER', '180');

INSERT INTO config_table (jobname, key, value) VALUES ('cake_api', 'POSTS_LIMIT', '50');

INSERT INTO config_table (jobname, key, value) VALUES ('cake_api', 'OFFSET', '1');

INSERT INTO config_table (jobname, key, value) VALUES ('cake_api', 'REPORT_ID', '66');

INSERT INTO config_table (jobname, key, value) VALUES ('cake_api', 'REPORT_VIEW_ID', '66');

INSERT INTO config_table (jobname, key, value) VALUES ('cake_api', 'CONTENT_TYPE', 'application/x-www-form-urlencoded; charset=UTF-8');

INSERT INTO config_table (jobname, key, value) VALUES ('cake_api', 'COOKIE', '<Paste cookie text here>');

INSERT INTO config_table (jobname, key, value) VALUES ('cake_api', 'API_KEY',
'3ODqKwgJVR3ufgDa9VInTSRZiHffqy');

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'LOG_DIR_PATH', '/home/ubuntu/logs');

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'TARGET_DIR_PATH', '/home/ubuntu/target');

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'STAGING_DIR_PATH', '/home/ubuntu/stage');

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'SMTP_PORT', '<Please provide SMTP port here>');

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'SMTP_HOST', '<Please provide SMTP host here>');

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'USER_NAME', '<Please provide SMTP username here>');

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'USER_PASSWORD', '<Please provide SMTP password here>');
```

- e. Run/Schedule script MASTER\_CAKE/ master\_cake\_run.sh (.bat file in same folder for windows server) . The script file should have read and execute privileges set properly for schedule to work.
- f. To check log go to \$LOG\_DIR\_PATH/cake\_api\_<current\_timestamp>.txt
- g. For every run an entry will be made to table runlog which will have start and end timestamp for run.

### **Output data tables**

Table name	Description
ck_leads_by_affiliates	Table containing leads by affiliates data
ck_leads_by_buyer	Table containing leads by buyer data
ck_posts	Table containing cake report post data

## VI. Deliverable name: MASTER\_BING\_CONFIG.zip(Bing configuration data processing batch)

Description: This archive contains script which will fetch from bing api for all accounts. The script name is MASTER\_BING\_CONFIG\_run.sh

Script dependencies – Script requires config\_table to be loaded with required data. Also configuration file containing database parameter values should be present at following path: /home/ubuntu/config\_files/bing\_config.txt

Follow following steps for installation:

- a. Create an entry in table jobs such as:

```
INSERT INTO jobs ("FromEmailID", "JobName", "JobDescription", "JobSuccessEmailID", "JobFailureEmailID", "CreatedDate", "UpdatedDate", "IsActive", "JobSuccess_CC_EmailID", "JobFailure_CC_EmailID", "JobGroup", "JobID") VALUES ('cybytetech@yahoo.com', 'MASTER_BING_CONFIG', NULL, 'poonamt@cybytetech.com', 'poonamt@cybytetech.com', NULL, NULL, 1, 'poonamt@cybytetech.com', 'poonamt@cybytetech.com', NULL, 6);
```

**Note: Please put required email addresses where you want respective email notification to be delivered. If you want success email to be delivered to multiple email accounts, separate them by a comma. Likewise for other email lists.**

- b. Unzip MASTER\_BING\_CONFIG.zip on server.
- c. Please make sure cake\_config.txt at following path: /home/ubuntu/config\_files

Also config files should have following entries:

```
DB_HOST=<Database host name>
DB_NAME=<Database name>
DB_PORT=<Database Port>
DB_SCHEMA=<Database Schema>
DB_USER=<Database User>
DB_PASSWORD=<Database Password>
```

- d. Please make sure following entries are present in table config\_table:

```
INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'BING_GRANT_CODE', <Put client id here>);
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'REPORTS_DIR', '/home/ubuntu/stage');
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', <Place for phantom js browser on drive>);
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'BING_DEV_TOKEN', <put dev token here>);
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'REPORT_ID', '66');
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'APP_TYPE', 'desktop');
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'STAGING_DIR_PATH', '/home/ubuntu/stage');
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'BING_CUSTOMER_ID', '13055485');
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'BING_USER_ID', <Put user id here>);
```

```
INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'BING_PASSWORD', <Put password here for account>);
```

```

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'LOG_DIR_PATH', '/home/ubuntu/logs');

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'TARGET_DIR_PATH', '/home/ubuntu/target');

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'STAGING_DIR_PATH', '/home/ubuntu/stage');

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'SMTP_PORT', <Please provide SMTP port here>);

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'SMTP_HOST', <Please provide SMTP host here>);

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'USER_NAME', <Please provide SMTP username here>);

INSERT INTO config_table(jobname, key, value) VALUES ('common', 'USER_PASSWORD', <Please provide SMTP password here>);

```

- e. Run/Schedule script MASTER\_BING\_CONFIG/MASTER\_BING\_CONFIG\_run.sh (.bat file in same folder for windows server) . The script file should have read and execute privileges set properly for schedule to work.
- f. To check log go to \$LOG\_DIR\_PATH/bing\_config\_data\_<current\_timestamp>.txt
- g. For every run an entry will be made to table runlog which will have start and end timestamp for run.

### **Output data tables**

<b>Table name</b>	<b>Description</b>
bing_campaigns	Table containing bing campaigns data
bing_ad_grps	Table containing bing ad groups data
bing_ads	Table containing bing ads data

## **VII. Deliverable name: MASTER\_BING\_PERFORMANCE.zip(Bing performance data processing batch)**

**Description:** This archive contains script which will fetch from bing api for all accounts. The script name is MASTER\_BING\_PERFORMANCE\_run.sh

Script dependancies – Script requires config\_table to be loaded with required data. Also configuration file containing database parameter values should be present at following path: /home/ubuntu/config\_files/bing\_config.txt

Follow following steps for installation:

- a. Create an entry in table jobs such as:

```

INSERT INTO jobs ("FromEmailID", "JobName", "JobDescription", "JobSuccessEmailID", "JobFailureEmailID", "CreatedDate", "UpdatedDate", "IsActive", "JobSuccess_CC_EmailID", "JobFailure_CC_EmailID", "JobGroup", "JobID") VALUES ('cybytetech@yahoo.com', 'MASTER_BING_PERFORMANCE', NULL, 'poonamt@cybytetech.com', 'poonamt@cybytetech.com', NULL, NULL, 1, 'poonamt@cybytetech.com', 'poonamt@cybytetech.com', NULL, 6);

```

**Note: Please put required email addresses where you want respective email notification to be delivered. If you want success email to be delivered to multiple email accounts, separate them by a comma. Likewise for other email lists.**

- b. Unzip MASTER\_BING\_PERFORMANCE.zip on server.
- c. Please make sure `bing_config.txt` at following path: `/home/ubuntu/config_files`

Also config files should have following entries:

```
DB_HOST=<Database host name>
DB_NAME=<Database name>
DB_PORT=<Database Port>
DB_SCHEMA=<Database Schema>
DB_USER=<Database User>
DB_PASSWORD=<Database Password>
```

- d. Please make sure following entries are present in table `config_table`:

```
INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'BING_GRANT_CODE', <Put client id here>);

INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'REPORTS_DIR', '/home/ubuntu/stage');

INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', <Place for phantom js browser on drive>);

INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'BING_DEV_TOKEN', <Put dev token here>);

INSERT INTO config_table (jobname, key, value) VALUES ('cake_api', 'REPORT_ID', '66');

INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'TIME_PERIOD', 'YESTERDAY'); --Change it to
LAST_THREE_MONTHS for 3 months, if required

INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'APP_TYPE', 'desktop');

INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'STAGING_DIR_PATH', '/home/ubuntu/stage');

INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'BING_CUSTOMER_ID', '13055485');

INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'BING_USER_ID', <put user id>);

INSERT INTO config_table (jobname, key, value) VALUES ('bing_api', 'BING_PASSWORD', <Put password here>);

INSERT INTO config_table (jobname, key, value) VALUES ('common', 'LOG_DIR_PATH', '/home/ubuntu/logs');

INSERT INTO config_table (jobname, key, value) VALUES ('common', 'TARGET_DIR_PATH', '/home/ubuntu/target');

INSERT INTO config_table (jobname, key, value) VALUES ('common', 'STAGING_DIR_PATH', '/home/ubuntu/stage');

INSERT INTO config_table (jobname, key, value) VALUES ('common', 'SMTP_PORT', <Please provide SMTP port
here>);

INSERT INTO config_table (jobname, key, value) VALUES ('common', 'SMTP_HOST', <Please provide SMTP host
here>);

INSERT INTO config_table (jobname, key, value) VALUES ('common', 'USER_NAME', <Please provide SMTP username
here>);

INSERT INTO config_table (jobname, key, value) VALUES ('common', 'USER_PASSWORD', <Please provide SMTP
password here>);
```



- e. Run/Schedule script MASTER\_BING\_PERFORMANCE/MASTER\_BING\_PERFORMANCE\_run.sh (.bat file in same folder for windows server) . The script file should have read and execute privileges set properly for schedule to work.
- f. To check log go to \$LOG\_DIR\_PATH/bing\_api\_<current\_timestamp>.txt
- g. For every run an entry will be made to table runlog which will have start and end timestamp for run.

### **Output data tables**

<b>Table name</b>	<b>Description</b>
bing_ad_performance_report	Table containing bing ad performance report data
bing_keyword_performance_report	Table containing bing keyword performance report data
bing_search_query_performance	Table containing bing search query performance data

## **Required setup files**

Following files are required for some sources:

### **Adwords key file:**

Put adwords key file for adwords in a folder on which server is running. The key file has been placed on VM at following path

/home/ubuntu/deliverables/setup\_files/test-3839ced14898.p12

Please make sure the full path to file(with filename as well) is being set in config\_table where key='KEY\_FILE' and jobname='api\_adwords'. This file is required to authorize data fetch from adwords

### **Bing Phantom JS browser file:**

Download Phantom JS Browser from following path on web:

[https://bitbucket.org/ariya/phantomjs/downloads/phantomjs-2.1.1-linux-x86\\_64.tar.bz2](https://bitbucket.org/ariya/phantomjs/downloads/phantomjs-2.1.1-linux-x86_64.tar.bz2)

Unarchive above file. Please make sure the full path to browser(file located in bin folder) is set into config\_table where jobname='bing\_api' and key='PHATOM\_JS\_PATH'

Path will look like;

e.g.

<Path where unarchived above archive> /phantomjs-2.1.1-linux-x86\_64/bin/phantomjs

### **Configuration files for all the connectors**

For all the connectors, configuration files are required to provide database parameters. The files are uploaded at test VM at following path:

/home/ubuntu/deliverables/config\_files

Added are the dummy values. One need to change them to actual database values. Also please make sure while installation all the configuration files are placed at following path on server:

/home/ubuntu/config\_files

## **Required setup tables**

Following setup tables are required for scripts to run:

- Schema and data for jobs table – This table contains entry for all ETL jobs and contains flag “IsActive” which can decide whether job needs to be run or not.  
If this flag is set to 1, job will be run by scheduler otherwise not.
- Schema for runlog table – This table will contain one entry for each run for each job.
- Schema and data for adwords\_clients – This table contains adwords clients and active flag which decide for which all clients data is to be fetched. Data will only be fetched for client for whom active flag is TRUE.
- Schema and data for config\_table – This table contains configuration data for all ETL jobs. Please make sure the script for same is run with environment specific values before running any ETL process.

DB scripts for all of the above has been placed on test VM at path

/home/ubuntu/deliverables/db\_scripts